2

HDL-PR-91-1

July 1991

A Survey of Techniques in Aided Target Recognition (ATR) from Digital and Optical Perspectives

by Charles G. Garvin

U.S. Army Laboratory Command
Harry Diamond Laboratories
Adelphi, MD 20783-1197

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1991	Progress, 6/88 to 4/89

4. TITLE AND SUBTITLE
A Survey of Techniques in Aided Target Recognition (ATR) from Digital and Optical Perspectives

5. FUNDING NUMBERS
PE: 62709A

6. AUTHOR(S)
Charles G. Garvin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Harry Diamond Laboratories
2800 Powder Mill Road
Adelphi, MD 20783-1197

8. PERFORMING ORGANIZATION REPORT NUMBER
HDL-PR-91-1

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Laboratory Command
2800 Powder Mill Road
Adelphi, MD 20783-1145

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
AMS code: 612709.H950011
HDL PR: DE18D1

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A survey of a number of useful linear and nonlinear image processing techniques for aided target recognition tasks is presented, and a summary of the current progress with both digital and optical implementations of these techniques is given.

14. SUBJECT TERMS
ATR, aided target recognition, processing, signal processing

15. NUMBER OF PAGES
26

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	17. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

Contents

Illustrations

Executive Summary

There is no generally applicable definition of the term *aided target recognition* (ATR). It has as many definitions as there are target-recognition tasks. I take as a working definition the task of mapping scenes to representations and extracting information concerning specific elements from these representations, such as the physical attributes of objects in the scenes.

The overall problem area serves as an umbrella for work in many fields, including optical processing and both analog and digital electronic processing. Because work in these fields mostly addresses specific processing functions and focuses on performing these functions optimally, this work does not often generalize. If the general ATR problems can be well formulated, it may be possible to direct work in these fields toward promising areas.

Several important questions are open that are fundamental to ATR problems, such as optimal data representation, image-from-scene mapping, preattentive/attentive vision boundaries, and the separability of the variables in a model. The bulk of current ATR work makes assumptions about these (and other) questions which, if mistaken, could render this work invalid. For example, if it is possible to show that information from shift-invariant processing, from scale-invariant processing, and from rotationally invariant processing combined is equivalent to information obtained using a processing scheme invariant to all three simultaneously, then the information is separable; if not, information is lost. Unless a proof of such separability can be given, it is more prudent to assume that the information is not separable.

A great many image sensors are optical, but the information is almost always converted into analog or digital electronic signals by the sensor. A step can be added to perform analog optical preprocessing of the image information. Tasks such as cueing, filtering, or data reduction could be accomplished in an optical preprocessing stage. For example, many researchers believe that edge enhancement is an essential operation for pattern recognition; it is possible to perform Sobel edge enhancement using liquid-crystal spatial light modulators (SLM's) as a preprocessing step to image detection. In edge enhancement, data are reduced and a preferred data representation is selected. This sort of preprocessing step exploits the capabilities of optics, such as large information throughput rate, continuous mapping, and parallel noninterfering connection at an appropriate place in the data pipeline; issues of programmability and calculation accuracy place constraints on such preprocessors.

Based on an evaluation of existing algorithms and devices, it is possible to make a few forecasts on the success of ATR systems. At this time it is possible to use ATR systems to do on-line product inspection for a small number (say 10) of well-known defects in parts on an assembly line. This assumes a simple geometric object and a processor capable of doing real-time recognition with either shift, scale, or rotational invariance, but not all three at once. Within five years it should be possible to perform the same task with shift-scale-rotation invariance in a lab environment; moreover, hybrid electronic/optical systems should be available to do some simple recognition tasks on airborne platforms, such as moving target indication, novelty filtering, or tracking for a small number of well-known targets. Systems that can perform emergent feature recognition (for recognizing unknown targets) on images with clutter (that is, real-world images) will require advances in algorithms and devices that should take at least 10 years to develop at the present rate of progress.

It is possible to evaluate the state of the available technologies in terms of a typical goal: to fly a smart munition capable of detecting, classifying, tracking, and targeting an enemy asset in the presence of clutter in the air, on sea, or on land. The devices that are currently used fall into three areas: minibench optics, integrated optics, and digital electronics.

It is already possible to construct minibench optics to do the above task, but the architectures and components do not yet exist to do the detection, recognition, tracking, and targeting to useful limits.

The development of integrated optics is even less far along: there are no existing system-level devices, since at this point research is active in designing components necessary to build systems such as waveguides, modulators, combiners, lenses, and detectors. No common material (such as lithium niobate or gallium arsenide) has yet been identified for monolithic structures.

Digital electronic processing is in the favorable position of having architectures and system-level components (as well as research-level VHSIC components), but none of those available are fast enough to perform the above tasks to acceptable levels. Backward compatibility and programmability can be considered additional advantages of digital electronic processing, but both of these advantages are often traded away in an attempt to increase processing speed.

1. Introduction

Although there are many possible definitions of the term *aided target recognition* (ATR), I take as a working definition the task of mapping scenes to representations and extracting information concerning specific elements from these representations, such as the physical attributes of objects in the scenes.

The methods of attacking problems in ATR can be separated (for this discussion) into several competing strategies.

(1) The most basic image manipulation, parameter extraction, is the most commonly used technique in ATR applications.

(2) The next logical grouping of image manipulation techniques involves the first level of mathematical abstraction of images using linear transform techniques. This grouping includes linear decomposition techniques as well as integral transform techniques. Matched filtering, correlation, and Fourier decomposition are some of the common examples of linear transform techniques.

(3) Nonlinear techniques applied to ATR make up a third group. This grouping includes neural network techniques (which I choose to call nonlinear transform techniques) to ATR problems.

This paper examines the competing methods used for ATR in terms of complexity of implementation, calculation burden, and operational robustness. It will be necessary to consider data compression techniques and the influence they have (if applicable) on competing methods. The most serious challenges to any ATR solution are perception invariance (invariance with respect to transformations: i.e., stimulus equivalence) and image generalization or abstraction (segmentation: i.e., feature extraction). These must be considered in examining ATR techniques.

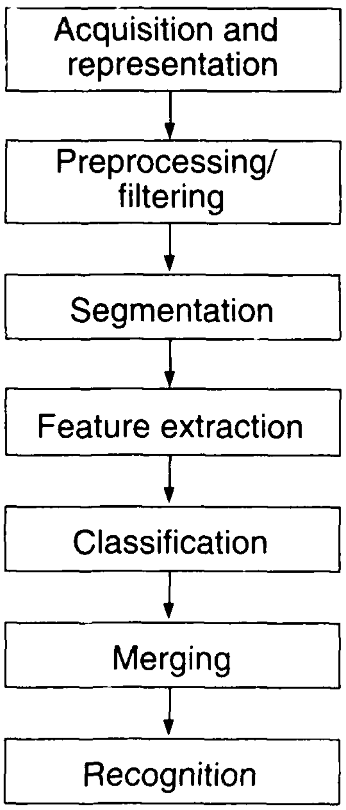
2. Survey of ATR Techniques

This survey is intended to present some of the most commonly used techniques for attacking ATR problems in a common format, comparing the order of abstraction from raw image data and the complexity of calculation.

Can one come up with a canonical sequence for ATR processing? No such sequence applies equally well to all analog and digital ATR mechanisms, but I use the sequence in figure 1 [1] as a point of departure.

Target data are acquired using any number of sensors operating alone or in data fusion, and the data are represented by some means as a temporally and/or spatially varying signal or image. I shall assume that we either detect or construct an image from these data. The image formed can be segmented into subpictures; at this point some subpictures may be determined to be uninteresting and can be eliminated. At any point in this sequence, of course, redundant data can be filtered at the cost of increasing processing time and complexity. After segmentation, the data in the subpictures are filtered for features. The linear and nonlinear processing algorithms listed below are some mechanisms used for this processing step. Features extracted by these

Figure 1. Sequence of processing of ATR data (adapted from Hoffman and Jain [1]).



mechanisms can be stored (if simple data collection is the goal), or compared with stored feature information (either previously collected or model generated) to complete the identification process. Often image analysis routines are divided into preattentive (or early warning) vision and attentive vision problems. Algorithms which operate in either area should be fast, but for preattentive vision problems speed is critical.

2.1 Parameter Extraction

Some of the simplest and most powerful picture-analysis methods apply only first-moment techniques. Much can be determined from simple first moments such as pixel intensity distributions, spatial frequency distributions, or temporal frequency distributions. Rigorously defined models that describe the first-moment characteristics of image data offer a means to make first-moment techniques more robust. The use of distribution theory to perform the identification process provides another means of adding mathematical meat to the skeleton of first-moment analysis.

Some model-based systems are designed to predict measurable quantities such as the pixel intensity distribution (possibly parameterized by aspect angle) observed in an image. These systems use models of objects, clutter, channel noise, etc, to generate such predictions. Observed data are compared with these calculations, and the results are used to identify objects. In general, the image representation models are simplified; most often the images are modeled by vectors whose elements correspond to pixel grey levels or other image parameters (such as edges, zero crossings, etc).

A model-based scene-representation method [2] known as "maximum a posteriori" (MAP) estimation takes the model-based technique a little further. This is one example of the many methods of ascertaining the optimal model for mapping a given scene to image data.

In these discussions, image data are seen as a result of the operation of some mapping A on real-world scenes:

$$i = A[s] \ .$$

The scene-to-image mapping A is in general nonlinear. The image-to-scene inverse mapping is then the fundamental aspect of the target recognition, and in this view the determination of A^{-1} is how the ATR problem is formulated. In general, there may not be a unique scene s that satisfies the equation

$$s = A^{-1}[i] \ .$$

Based on some assumptions about the statistical distributions of the image data and the noise, MAP constructs a cost function relating the image data, the unknown mapping A^{-1}, and any constraints on the system (such as continuity or bounds on image parameters). The cost function is optimized over the possible mappings, and an estimate for the scene is calculated based on the chosen A^{-1}. Similarities exist between MAP and the work of S. Geman and D. Geman [3] on so-called "stochastic annealing," which treats the optimization methods used to determine such mappings.

Local operators (such as gradient or Sobel) or global operators (such as integral transforms) are used in first-level filtering of image data for feature extraction. Whether it is called representation, decomposition, filtering, or correlation, this first data-reduction step compresses the image data into features that will be measured against stored signature data or used to construct stored signatures. As data compression (while retaining the significant data) is critical to any rapid ATR process, raw images are seldom examined in any sophisticated way in real-time applications. In image sequence analysis, stationary information may be subtracted away by a running image subtraction, leaving only image data that change at a fixed rate. Images are often filtered to emphasize high-frequency information (edge enhancement by Sobel or other local operators). The implementation of such an operator is often a correlation filter with a small kernel, such as a 3×3 pixel matrix.

2.2 Linear Processing Algorithms

The end goal of transform methods is to represent the total bulk of data to be processed without losing important information, ideally so that the data are reduced to a signature that is unique (orthogonal to all other signatures) and can be used to identify a target. Any complete set of linearly independent functions will do for the task of data compression (of signals or images) into coefficients. The differences among such functions can be evaluated in terms of their invariance properties, their fidelity of representation, and the ease and speed of their implementation [4–7].

2.2.1 *Hotelling or Eigenvector Decomposition*

In transform encoding a picture, the intent is to separate all the data into a set of independent points in a transform space so that they can be distinguished from one another. The closest we can come to such a transformation is the Hotelling transform, which produces uncorrelated but not necessarily independent representations. Given a picture with $N \times N$ pixels, each of which can take on 2^k grey-scale values, there would be $2^{N \times N \times k}$ possible points that represent pictures in an $N \times N$ space. To establish a coordinate system in which all these points are independent, the transformation from picture coordinates x_i (where i runs from 1 to N^2) to a new coordinate system y_i is an $N \times N$ dimensional rotation matrix A, so that

$$y_i = \sum_{0}^{N \times N} (A_{ij} x_j) \ .$$

Representing a given picture in terms of the new coordinates uses the inverse rotation A^{-1}, so that any given picture coordinate can be represented as

$$x_i = \sum_{j=1}^{N \times N} \left(A_{ij}^{-1} y_j \right) \ .$$

The A matrix in the Hotelling transformation is formed using the eigenvalues of the covariance matrix as the diagonal elements of an N^2 by N^2 matrix. Because of the necessity of performing an N^2 by N^2 matrix inversion in order to calculate the elements of the A matrix for the Hotelling transform, its implementation (analog or digital) is bound to be more complicated than other transformations. For this reason, although it yields the least mean-square error in image representation, it is rarely used to represent images when rapid processing is desired. Instead, one of several other transformation kernels is used; a few of these follow. In the following transform discussions, u and v are transform variables, while x, y, and t are signal/image variables. In the following discrete forms of the transforms, N refers to number of variables, unless otherwise indicated.

2.2.2 *Fourier Decomposition*

In Fourier decomposition, the orthogonal polynomials of the transformation are the sine and cosine functions. It appears that Fourier decomposition is the most used transform technique because of the ease of performing the necessary computation of coefficients by either analog or digital means and the adequacy of its asymptotic convergence to the eigenvector transform performance in mean square error. The familiar Fourier transform (FT) kernel is

$$A_{ut} = \frac{1}{N} \exp -i(2\pi ut) \ .$$

A great many ATR mechanisms (both digital and analog) rely on the Fourier representation of image data; examples include the Georgia Institute of Technology Research Institute's digital stationary and moving target recognition algorithms, Grossberg's adaptive resonance theory digital pattern-recognition programs, and optical FT holographic element correlators and spectrum analyzers.

2.2.3 *Walsh-Hadamard Decomposition*

In Walsh-Hadamard decomposition, the transform functions take on only the values 0 or 1, which simplifies digital computation tremendously. The kernel for a Walsh-Hadamard transform of order $N = 2^n$ is given by

$$A_{ut} = \frac{1}{\sqrt{N}} (-1) \exp\left[\sum_{k=0}^{n-1} \{b_k(u) b_k(t)\} \right] ,$$

where b_0 to b_{k-1} are the bits in the binary representation of the signal data.

2.2.4 *Discrete Cosine Decomposition*

The discrete cosine transform (DCT) uses the set of orthogonal polynomials known as the Chebyshev polynomials. Its popularity stems from the relative ease of digitally computing the transform, combined with a lower mean square error than is obtained with the Walsh-Hadamard transform or the discrete Fourier transform algorithms. The kernel for the DCT is

$$A_{ut} = \frac{2}{\sqrt{N}} K(i) \cos\left[(2u + 1)t \frac{\pi}{2N}\right] ,$$

where

$$K(i) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } i = 1 \\ 1 & \text{for } i = 2, 3, \ldots, N \\ 0 & \text{elsewhere} \end{cases} .$$

2.2.5 Gabor Decomposition [8]

In the above transform techniques, their one- or two-dimensional nature is merely a matter of notation, since all the transform variables are essentially of the same character. The Gabor transform differs in that it is a two-dimensional transform technique at the minimum. The kernel of the Gabor transform uses both time and frequency (or temporal and spatial) variables to represent the encoded image. The Gabor transform differs in another way from those above in that the transform basis vectors are not orthogonal (hence the coefficients are highly correlated with one another), and the transform is not reversible in the same way as those above.

Once again, since it requires significant computation time, the Gabor transform is not competitive when rapid processing is required. The Gabor kernel is

$$A_{uvxy} = \exp\left\{-\left[\frac{x^2}{2}(\Delta W)^2 + \frac{y^2}{2}(\Delta L)^2\right]\right\}\cos(2\pi + \theta) \ ,$$

where ΔW is the spatial width of the sensor and ΔL is the spatial height of the sensor; the resolution size of the data can be represented by these delta variables. Some ATR systems whose goal is to emulate biological vision systems use the Gabor representation for image data.

2.2.6 *Correlation Transform Techniques*

Correlation transformation can be looked at as a decomposition process, as can the transform techniques listed above. In correlation, the kernel used to decompose the data is data itself. For an autocorrelation, the data act as both kernel and data; in cross-correlation the data are compressed against a kernel that is stored or selected data. A typical scenario for image correlation would use a training set of various versions of the image of interest (possibly the target at various aspect angles, and/or multiple target images). These images would be correlated against one another and the correlation coefficients compressed to form a sort of proto-image. This would be correlated against test set(s), and recognition would consist of the result of this correlation exceeding some threshold.

Instead of the training set consisting of images, a set of action principles could be used for the construction of images. In this case, actions taken in building the image or reading the image are encoded, and a test image is compared against these action principles. If rebuilding or decomposing the image reveals similar or identical action principles, then recognition occurs. Many model-based systems take this approach.

2.3 Nonlinear Processing

Neural network (NN) solutions have been called "nonalgorithmic" because these systems are self-organized, rather than being programmed. All NN's consist of two or more layers of simple processing elements (PE's), which generally consist of a summation element and a thresholding function.

All NN approaches use some nonlinear element; at minimum, the NN uses a threshold step (most often a sigmoid function) as the nonlinear element. Thus NN techniques are classified as nonlinear processing methods because of this thresholding nonlinearity.

Each PE is connected to other PE's by a channel whose strength is modifiable through one of several learning laws (Hebbian, Grossberg, Hopfield, others) based on feedback from other layers in the NN. Training data (whose output result is known) are fed into the NN and passed through, and the NN output is compared with the desired output. Error signals based on this comparison feed back through the net via the learning rule to modify internal channel connections.

Most pattern recognizers using associative memory are neural nets; a possible way of looking at them is to view them as multiple-pass correlators having sets of filters or masks with which they have been trained, which select the closest filter or mask to any given input test image by performing a series of thresholded correlation calculations.

Other more sophisticated NN's are being examined and constructed to perform ATR tasks other than associative recall, such as data reduction, edge enhancement, or image segmentation.

3. Review of Experimental Progress

Implementation methods such as those mentioned above are invariably influenced by device issues. Implementation of digital algorithms in software on Von Neumann machines, or in specially constructed hardware, faces issues like the number of necessary computations, analog-to-digital and digital-to-analog converter clock speeds and bit depths, chip count, hardware size, and power. Optical analog (and digital) techniques face the limitations of the illumination, modulation, and detection devices used to construct the architectures mentioned above. For optical associative memory, architectures of choice have used magneto-optic spatial light modulators (MOSLM's), photorefractive media, and/or holographic media; relevant device capacity issues govern the size and memory depths of implemented systems, obtainable dynamic ranges, and resolution, as well as the choice of algorithm. Table 1 summarizes one of the digital implementation considerations (the computational burden) for a few of the integral transform techniques mentioned above.

Because many of the methods mentioned above have been worked on for decades, we concentrate on mentioning a few recent experiments which typify the areas.

3.1 Digital Implementations

Software image-processing programs on general-purpose digital computing machines, as well as specially constructed or "hardwired" digital circuitry, are frequently used for ATR applications. Thousands of specific processing problems fall into the ATR arena, and it seems that each problem has a specific method of digital solution. The abundance of different types of computing machines and programming languages combines with the inexhaustible array of image-processing algorithms to produce an impressive background of digital processing methods. In addition, the proliferation of array process-

Table 1. Computational burden for selected transform techniques

Transform	No. of arithmetic operations required	
	Real	Complex
Walsh-Hadamard	$N \log_2 N$ (additions or subtractions only)	—
Discrete Fourier	—	$N \log_2 N$
Discrete cosine	$N \log_2 N$	—
Hotelling/ Karhunen-Loeve/ eigenvector	N^2	N^2

ing hardware leads researchers to try larger and more complicated algorithms to solve their processing problems within the very real time constraints. In the sections below only a few typical digital electronic computing approaches to ATR problems are discussed.

3.1.1 Linear Processing

Model-based system/parameter extraction. An example of a model-based technique using parameter extraction is discussed by Flachs et al [9]. Based on assumed distribution functions for reflected energy and noise sources, this technique uses image input from a sensor (such as a focal plane array) to create task-dependent metric(s) corresponding to the detectability of targets in a selected environment. Analysis of the distributions of the complexity metric(s), including new input image data, indicates the presence or absence of targets in a given scene to a chosen confidence level. In an example task, "cuer complexity," "segmentation complexity," and "classification complexity measure" are calculated. The values for these three metrics are compared to bounds or thresholds to indicate (1) that target/background separation can be done; (2) that simple image data (such as grey level) can be used to make the separation; and (3) whether or not multiple targets can be separated using the same image data. Experiments conducted on a digital computer using infrared (IR) image data show recognition of an armored personnel carrier in cluttered (natural environment) background.

Transform methods. A series of image recognition experiments were conducted using a transform method known as discrete rectangular wave transform (DRWT) [10]. In the experiments, Walsh-Hadamard-like functions (rectangular waveforms taking on values of only 0 or 1) were used in the DFT digital algorithms instead of the set of orthogonal functions. Image information (edge-only outlines of aircraft) was transformed using DFT and DRWT algorithms, low-frequency transform coefficients were retained, and images were classified with respect to distance between their transform coefficients and library features. The methods were tested for rotation invariance with and without Gaussian noise (the signal to noise ratio (SNR) varied in the experiments from 30 to 3 dB), and results indicated the superiority of DRWT. The results demonstrated that the DRWT technique correctly identified 3-dB SNR rotated images with an accuracy of 33 to 66 percent in multiple trials.

3.1.2 Nonlinear Processing

Carpenter and Grossberg [11] have been active in the application of NN pattern classification to ATR problems for several years, develop-ing over this period a formalism known as adaptive resonance theory (ART) and implementing it in software on digital computers in pattern recognizers called ART1 and ART2. A recent paper [11] is indicative of some of their recent results in vehicle identification. Using ART2 on a digital computer, correct classifications were made on multiple samples of IR imagery and range data (differing in scale, rotation angle, and position) of trucks taken in Gaussian noise on four different categories. Reported results were 80-percent correct identification in 10-percent noise with no false alarms.

Another NN approach to invariant pattern recognition involves the layering of several parallel slabs of fully connected adaline (adaptive linear elements) neural elements onto adaptive layers of neurons using a layer of fixed-weight "majority-vote-taking" elements. The number of parallel slabs of adalines necessary will depend on the degree of invariance desired. The first layer of adalines is trained to classify patterns regardless of position, scale, or rotation, but its output is unusable; the layers of adaptable elements are used to unscramble the first layer's output. In the experiments reported by Widrow et al [12], 25 slabs of 5×5 adalines fed a two-layer adaptive net. This design gave translation-invariant recognition of 36 patterns to better than 98-percent accuracy after about 1000 learning cycles.

3.2 Optical Implementations

The bulk of the optical implementations of ATR techniques seem to cluster in two areas: integral transform techniques (Fourier, Hough, Wigner) and associative memory architectures. The former are implemented in various different technologies, but the basic architecture is the same: the signal is optically transformed, transform coefficients are matched to a "coefficient bank" stored in optical memory (such as a spatial light modulator (SLM), transparency, or hologram), the signal is classified as one of these or an outlyer, and outlyer coefficients are possibly stored in the coefficient bank.

Associative memory architectures are implemented with either so-called "inner-product" or "outer-product" schemes; in these schemes, received image data are passed into a multidimensional processing system which has been "trained" by a predetermined sample set, and the associative memory kicks out an identification or relates the ambiguity as an error.

Optical pattern recognition using holographic spatial filtering is a type of transform method. The information content of the spatial filter in implementations is severely limited in terms of the spatial bandwidth which can be represented (linewidth per millimeter limits) and/or the intensity levels (dynamic range) which can be used to represent information. For most real-time applications, SLM's offer only binary representation of information over a 100×100 pixel 1-cm^2 area. Reducing the information content to be represented by the filter is an approach; however, the discarded information cannot be necessary for valid recognition.

3.2.1 Linear Processing

Orthogonal polynomial techniques. One of the simplest of optical transform methods to implement uses the Fourier kernel. In the example discussed by Sheng et al [13], the distribution of the Fourier spectrum is used to characterize an image using Fourier-Mellin descriptors (FMD's). These FMD's describe the intensities of the Fourier spectrum representation of the image and are automatically shift invariant; rotation and scale invariance in this representation requires an additional normalization of the FMD. Since the loss of phase information creates ambiguities and the FMD representation is not one-to-one, the class of recognizable objects must exclude these ambiguities. The architecture uses a standard two-dimensional optical spectrum analyzer and digitally calculated matched filters to perform object classification on optically processed images.

Another transform kernel frequently implemented optically is the Hough transform (HT). Experiments discussed by Casasent and Richards [14] compare recognition results of using both the HT and FT on identical image data in a product inspection application. Two optical HT architectures were examined, differing in the choice of transform variables (one Cartesian, the other polar coordinates). The architectures were implemented using liquid-crystal televisions (LCTV's) to input camera image data to the processor; the FT was implemented by simply imaging the LCTV through a spherical lens onto the detectors.

The experiments were designed to find defects in wire terminals, classifying the fault as either "splayed" or "smashed," based on transform representations (which were normalized composites of many examples of faulty terminals) of each class. Results showed the rotation invariance of the FT as its greatest strength; HT processing gave best discrimination among faults, as well as quantitative information on the magnitude of faults, although the HT was dependent on aspect angle. Neither implementation required scale invariance, since in the application the image scale would be fixed. Conclusions indicated that a combination of techniques would be necessary to solve the general problem with shift and rotation.

Matched filter and correlation techniques. Optical correlators have been in widespread use in signal-processing applications for several years. The feasibility of using optical correlators for image processing and target-recognition applications has been shown with numerous lab demonstrations; however, realistic application of optical correlation architectures to current ATR problems has not been possible because of the marginal performance of two-dimensional SLM's. The emergence of low-cost two-dimensional SLM's in the form of LCTV's has increased efforts in this area.

A typical demonstration of optical image correlation for an ATR problem is given by Chao and Liu [15]. Using an interferometric time-integrating correlator architecture with holographic correlation filters, Chao and Liu demonstrated tracking of three (overhead view) model vehicles at TV frame rate. A video image encoded on the LCTV is correlated with spatially separated holographic matched filter (HMF) references, using a specially constructed holographic lens to image the LCTV onto each of the HMF's. The typical problems associated with correlation techniques (such as undesired partial correlation among objects) and holographic matched filters (limits on resolution and number of reference images) place constraints on the extension of this sort of demonstration to an ATR system. The multiplicity of HMF's allowed good translation invariance for this image correlator; however, since the angle (or rotation) sensitivity of the HMF's was high, angular invariance would require the encoding of more rotated images on the HMF's, reducing the SNR performance because of limitations in the recording material.

3.2.2 Nonlinear Processing

An example of associative memory (or content-addressable memory) is discussed by Farhat et al [16]. The system uses a vector-matrix multiplier with a nonlinear iterative feedback stage to implement a fully connected, two-layer network. The network deals with binary input vectors and can represent a 4×8 image. In the optical implementation, an array of 32 light-emitting diodes (LED's) illuminates a 64×64 element fixed memory mask, and the throughput light is detected by an array of photodiodes. Electronic circuitry following the photodiodes performs the thresholding function and feeds the results back to the input array of LED's. Weight changes and learning can occur only if the fixed memory mask is changed or replaced by a modifiable spatial light modulator. Results demonstrated convergence to patterns having bit-error rates as high as 30 percent at cycle times of 60 ms.

There are several examples of optical holographic associative memory using photorefractive (PR) materials and phase conjugation [17,18]. Numerous experimental architectures exist; however, all are similar in their use of the photorefractive elements for gain and/or thresholding. In the work of Lee et al [18], the PR materials act as scratchpad storage media as well as gain and thresholding elements. The NN implemented is an inner-product matrix-vector multiplier; since the active elements are continuous media, the processor is effectively pixelized only by optical diffraction limits. Calculations indicate that it is possible to build a 23×23 neural element slab, fully connected, which can be cycled an arbitrary number of times because of the gain medium. In an experiment discussed by Lee et al [18], the system was trained on two 35-mm slides of an M-1 tank (overhead views). Results reported indicate a 3-s convergence time for a partially obscured and rotated (45° and 90°) image. The system reported is reconfigurable to act on other problems, such as novelty detection and parameter optimization, with commensurate resolution.

All existing biological neural systems use frequency and phase encoding and processing of information; however, all hardware implementations simulating such systems use amplitude and/or phase encoding and generally use fewer levels for number representation. These fundamental differences, along with the relatively simple learning rules and the small numbers of neurons and connections, contribute to the low processing capabilities of existing neural net systems.

4. Remarks

All the military services have programs concentrating on ATR. Although the requirements vary, the technology does not, so all ATR program officers get the same answers no matter which service they are in. Scenarios for the military use of ATR systems range from soldier-in-the-loop skill augmenters to completely autonomous robot weapons.

ATR is currently the problem generating the most noise at the most levels in the most services. Military aircraft pilots are overburdened with the tasks of operating the complex machinery necessary to accomplish their flight mission, and they do not have the resources to keep themselves and their aircraft alive while they are performing their mission. In the Army, tank commanders are faced with a similar problem on the ground, as are helicopter pilots in the air.

The same demand is heard from any close combat command: supply us with a device that targets enemy assets in real time. Weapons designers take up the chorus for targeting devices for their smart weapons. However, target-acquisition (and tracking) barriers limit the usefulness of smart munitions to situations in which target-acquisition and tracking problems are easily solved. Thus, surface-skimming cruise missiles are lethal at sea using passive radar homing devices, simply because their targets are so easy to see and track. Passive IR homing missiles are responsible for over 90 percent of the aircraft shot down for the same reason. This is why the close combat people complain about needing ATR—to use as counter-measures against these weapons.

The above sections survey only a few of the active research areas in ATR, but some conclusions can be drawn from the scope of the work that is seen and the types absent. In general, image data manipulation techniques are based on well-proved, rigorous mathematical theories that have been physically tested using all types of analog and digital computing systems. However, the choice of image acquisition and processing techniques is generally based on only heuristic arguments rather than being derived from first principles.

Although the so-called ATR problem has been studied for decades, it is only now that many researchers are attempting to solve the front end of the ATR problem, which some call preattentive vision and others call image-from-scene mapping. Because most of the experimental work mentioned above is application-driven, one can assume a form for image data; however, without a strong mathematical model for the formation of image data from scenes, diverse applications and the specific solutions to them cannot be compared across the board on a general metric to determine optimal solutions or preferred architectures.

References

1. R. Hoffman and A. Jain, *Segmentation and Classification of Range Images*, IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-9**, No. 5 (September 1987).

2. D. Kersten, A. J. O'Toole, M. E. Sereno, D. C. Knill, and J. A. Anderson, *Associative Learning of Scene Parameters from Images*, Appl. Opt. **26**, No. 23 (1 December 1987), 4999–5006.

3. S. Geman and D. Geman, *Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images*, IEEE Trans. Pattern Anal. Machine Intell. **PAMI-6**, No. 6 (November 1984), 721–741.

4. M. Kunt, *Redundancy Reduction in Digital Images*, Int. J. Opt. Electron. **3**, No. 1 (1988), 3–44.

5. P. Wintz, *Transform Picture Coding*, Proc. IEEE **60**, No. 7 (1972).

6. A. Netravali and J. Limb, *Picture Coding: A Review*, Proc. IEEE **68**, No. 3 (March 1980).

7. D. Elliott and K. Rao, *Fast Transforms: Algorithms, Analyses, Applications*, Academic Press, NY (1982).

8. D. Field, *Relations Between the Statistics of Natural Images and the Response Properties of Cortical Cells*, J. Opt. Soc. Am. A **4**, No. 12 (December 1982), 2379–2393.

9. G. Flachs, J. Jordan, and J. Carlson, *Summary of EVAL Activities Involving Multisensor Fusion Research for the Period March 1, 1987 through August 1, 1987*, contract DAAD07-83-C-0031.

10. O. Ersoy and D. Kim, *Image Recognition with the Discrete Rectangular Wave Transform*, J. Opt. Soc. Am. A **5**, No. 1 (January 1988).

11. G. Carpenter and S. Grossberg, *Invariant Pattern Recognition and Recall by an Attentive Self-Organizing ART Architecture in a Non-Stationary World*, Proc. ICNN (June 1987).

12. B. Widrow, R. Winter, and R. Baxter, *Layered Neural Nets for Pattern Recognition*, IEEE Trans. Acoust. Speech Signal Process. **ASSP-36**, No. 7 (July 1988).

13. Y. Sheng, C. Lejeune, and H. Arsenault, *Frequency-Domain Fourier-Mellin Descriptors for Invariant Pattern Recognition*, Opt. Eng. **27**, No. 5 (May 1986), 354–357.

14. D. Casasent and J. Richards, *Optical Hough and Fourier Transform Processors for Product Inspection*, Opt. Eng. **27**, No. 4 (April 1988).

15. T.-H. Chao and H.-K. Liu, *Real-Time Optical Holographic Tracking of Multiple Objects*, Appl. Opt., Jan 15, 1989, Vol 28, No. 2, p 226.

16. N. Farhat, D. Psaltis, A. Prata, and E. Paek, *Optical Implementation of the Hopfield Model*, Appl. Opt. **24**, No. 10 (May 1985).

17. H. Winston, *Optical Data Processing*, 1988 IR&D report, Hughes Aircraft Co.

18. L. S. Lee, H. M. Stoll, and M. C. Tackitt, *A Continuous-Time Optical Neural Network Associative Memory*, Optics Letter, Feb 1, 1989, Vol. 14, No. 3, p 162.

Distribution

Administrator
Defense Technical Information Center
Attn DTIC-DDA (2 copies)
Cameron Station, Building 5
Alexandria, VA 22304-6145

Director
Defense Communications Agency
Command & Control Center
Attn Tech Library
Washington, DC 20305

Director
Defense Communications Engineering
Center
Attn Code R123, Technical Library
1860 Wiehle Ave
Reston, VA 22090

Director
Defense Nuclear Agency
Attn Tech Library
Washington, DC 20305

Under Secretary of Defense for Research
& Engineering
Attn Research & Advanced Tech
Washington, DC 20301

US Army ETDL
Attn SLCET-DD
Ft Monmouth, NJ 07703

Commander
US Army Communications–Electronics
Command
Attn R&D Technical Library
Ft Monmouth, NJ 07703

Commander
US Army Computer Systems Command
Attn Technical Library
Ft Belvoir, VA 22060

Chief
US Army Research Office (Durham)
Attn SLCRO-MA, Dir Mathematics Div
PO Box 12211
Research Triangle Park, NC 27709-2211

Dept of the Air Force, HQ
6585th Test Group (AFSC)
Radar Target Scatter Facility
Attn LTC R. L. Kercher, Chief
Holloman AFB, NM 88330

Director
NASA, Goddard Space Flight Center
Attn 250, Tech Info Div
Greenbelt, MD 20771

NASA Langley Research Center
Attn Code 2620, Tech Library Br
Hampton, VA 23665

NASA Lewis Research Center
Attn Library
21000 Brookpark Rd
Cleveland, OH 44135

Director
NASA
Attn Technical Library
John F. Kennedy Space Center, FL 32889

Engineering Societies Library
Attn Acquisitions Dept
345 E. 47th Street
New York, NY 10017

Institute for Telecommunications Sciences
Attn Library
Boulder, CO 80303

The American University
Dept of Mathematical Statistics
Attn S. D. Casey
4400 Massachusetts Ave, NW
Washington, DC 20016-8050

Distribution (cont'd)

George Mason University
Attn ECE Dept., R. A. Athale
Fairfax, VA 22030

University of Maryland
Mathematics Dept. & Systems Research Ctr
Attn C. A. Berenstein
College Park, MD 20742

US Army Laboratory Command
Attn AMSLC-DL, Director, Corporate Labs

Installation Support Activity
Attn SLCIS-CC, Legal Office

USAISC
Attn AMSLC-IM-VA, Admin Ser Br
Attn AMSLC-IM-VP, Tech Pub Br (2 copies)

Harry Diamond Laboratories
Attn Laboratory Directors
Attn SLCHD-TL, Library (3 copies)
Attn SLCHD-TL, Library (WRF)
Attn SLCHD-NW, Chief
Attn SLCHD-NW-CS, Chief
Attn SLCHD-NW-E, Chief

Harry Diamond Laboratories (cont'd)
Attn SLCHD-NW-EH, Chief
Attn SLCHD-NW-EP, Chief
Attn SLCHD-NW-ES, Chief
Attn SLCHD-NW-R, Chief
Attn SLCHD-NW-RP, Chief
Attn SLCHD-NW-RS, Chief
Attn SLCHD-NW-TN, Chief
Attn SLCHD-NW-TS, Chief
Attn SLCHD-ST-OP, Chief
Attn SLCHD-ST, R. Johnson
Attn SLCHD-ST-OP, A. Filipov
Attn SLCHD-ST-OP, J. Goff
Attn SLCHD-ST-OP, N. Gupta
Attn SLCHD-ST-OP, L. Harrison
Attn SLCHD-ST-OP, D. Kafig
Attn SLCHD-ST-OP, J. Mait
Attn SLCHD-ST-OP, D. McGuire
Attn SLCHD-ST-OP, D. Prather
Attn SLCHD-ST-OP, B. Sadler
Attn SLCHD-ST-OP, M. Taylor
Attn SLCHD-ST-SA, J. Dammann
Attn SLCHD-ST-SA, B. Stann
Attn SLCHD-ST-SS, D. Gerstman
Attn SLCHD-ST-SS, J. Griffin
Attn SLCHD-ST-SS, D. McCarthy
Attn SLCTO, N. Berg
Attn SLCTO, B. Weber
Attn SLCHD-ST-SS, C. Garvin (20 copies)